

A wide-angle photograph of a majestic mountain range under a blue sky with scattered white clouds. The mountains are rugged, with dark, rocky peaks and patches of white snow clinging to their slopes. In the foreground, two hikers are standing on a rocky trail, looking towards the vast expanse of mountains. One hiker is wearing a light blue shirt and black shorts, while the other is wearing a dark t-shirt and black shorts. They are both holding trekking poles.

USABILITY TESTING

Metodologia dei test

Scelta dei partecipanti

Per la scelta dei **6 partecipanti** ci siamo basati sui seguenti parametri:

- utenti target: devono essere possibili utilizzatori dell'app;
- estraneità al progetto: non devono avere conoscenze pregresse del nostro progetto;
- background informatico: devono avere un minimo di esperienza con le applicazioni mobile;
- diverificazione degli utenti: deve esserci almeno un utente lead.

Inoltre abbiamo inquadrato meglio ogni partecipante tramite un **questionario pre-test**, in modo da poter trovare eventuali correlazioni (es. esperienza-successo) durante l'analisi dei risultati.

Modalità di esecuzione

I test hanno avuto luogo in posti che fanno sentire il partecipante a suo agio, come l'abitazione stessa, o sedi da lui frequentate.

La durata di ogni test è stata inferiore a 15 minuti.

Per quanto riguarda l'attrezzatura:

- il partecipante è stato munito di un computer su cui ha effettuato i test da Figma;
- il facilitatore era munito di una copia cartacea dello script e di una penna, per far compilare i questionari al partecipante;
- l'osservatore era munito di un telefono, con cui si è occupato della registrazione audio-video del test, e di carta e penna, con cui ha annotato le sue osservazioni sul comportamento del partecipante.

La suddivisione delle coppie è stata la seguente: Ferrero-Gerges, Galli-Guarnerio, Nidasio-Russo. Ogni si è alternata nel ruolo di facilitatore e osservatore nei due test effettuati.

ad ogni partecipante è stato fornito il modulo per il consenso informato

Criteri di valutazione

Per ogni task abbiamo valutato la prestazione in base alle seguenti metriche:

- tempo di esecuzione del task rispetto al tempo previsto (calcolato da noi a seguito di prove interne);
- tasso di successo del task (complete success, success with a minor issue, success with a major issue, failure);
- numero di aiuti richiesti al facilitatore.

Inoltre abbiamo proposto all'utente dei metodi ufficiali di valutazione soggettiva dei singoli task (con **domande post-task**), e dell'applicazione in generale (con un **questionario post-test**).

Questionario pre-test

Questionario pre-test

Età: []

Occupazione:

- [] studente
- [] lavoratore
- [] studente lavoratore
- [] altro

Quanto spesso vivi esperienze nella natura:

- [] più volte a settimana
- [] una volta a settimana
- [] 1-2 volte al mese
- [] raramente
- [] mai

Che tipo di esperienze prediliги:

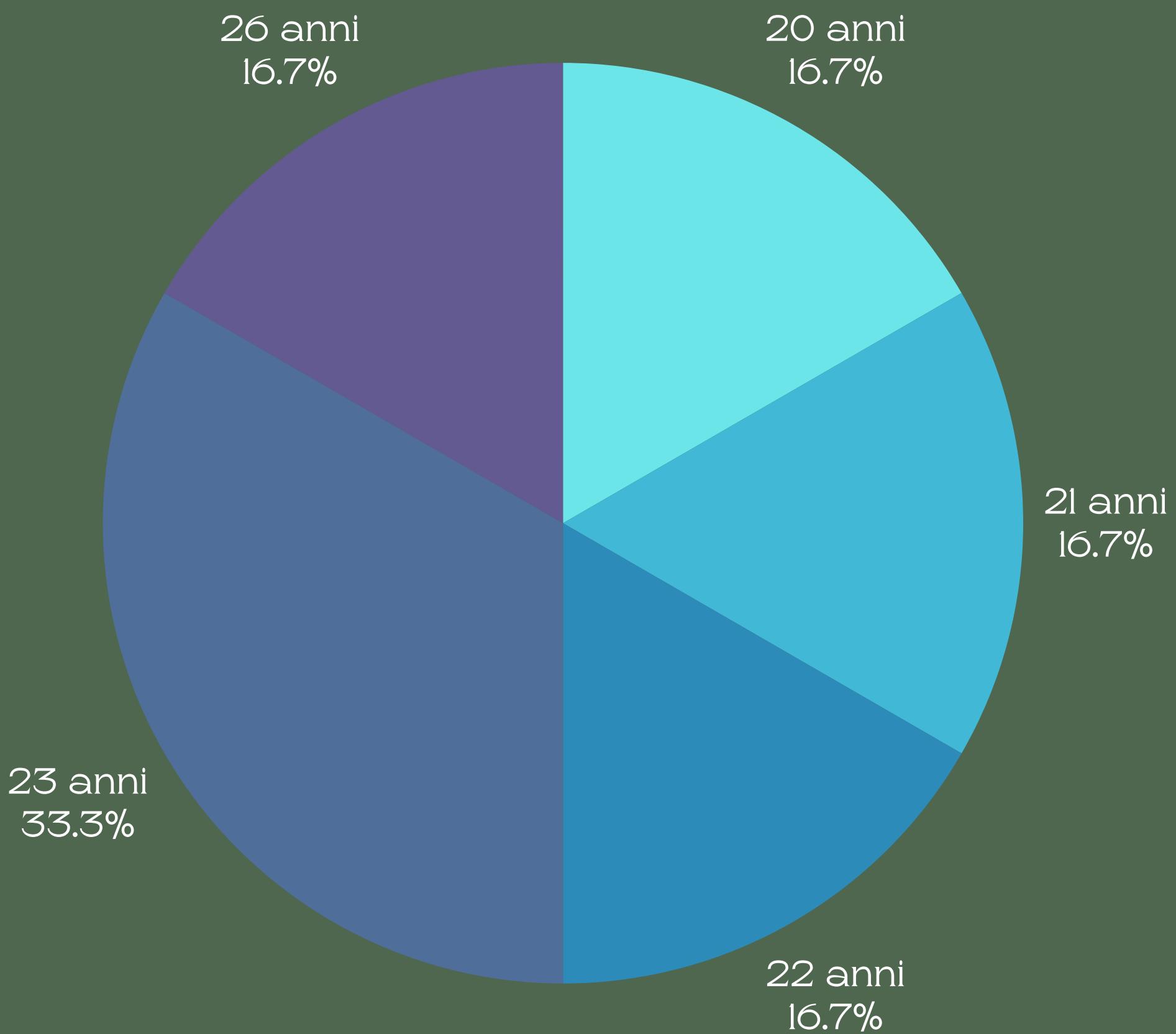
- [] trekking
- [] ciclismo
- [] sport acquatici
- [] sport estremi
- [] altro: _____

Quanta familiarità hai con app per organizzare eventi:

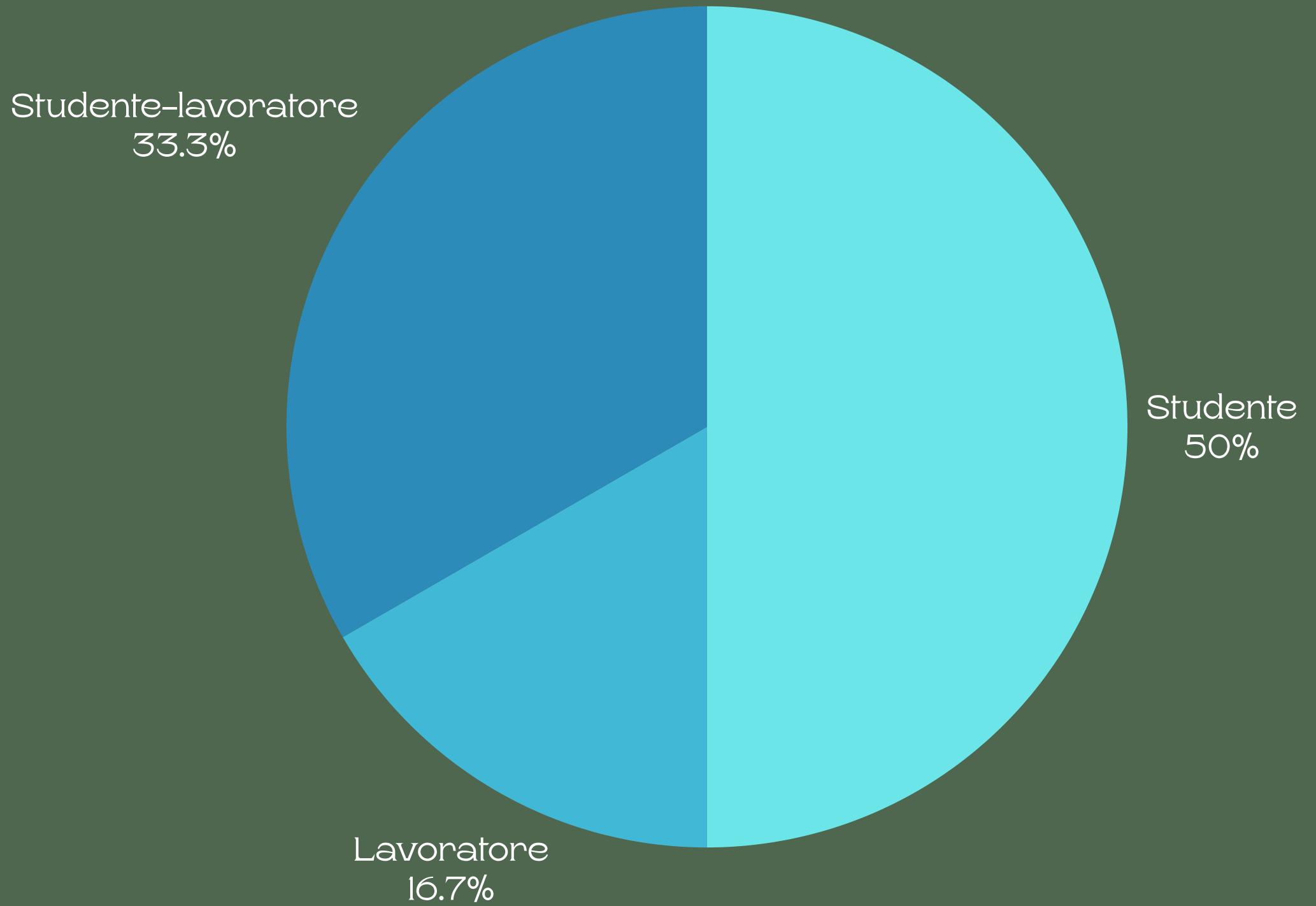
- [] mai usate
- [] le ho usate un paio di volte
- [] le uso frequentemente

Con questo questionario abbiamo analizzato l'anagrafica dei nostri partecipanti, la loro affinità con la natura, e la loro esperienza con strumenti simili alla nostra applicazione.

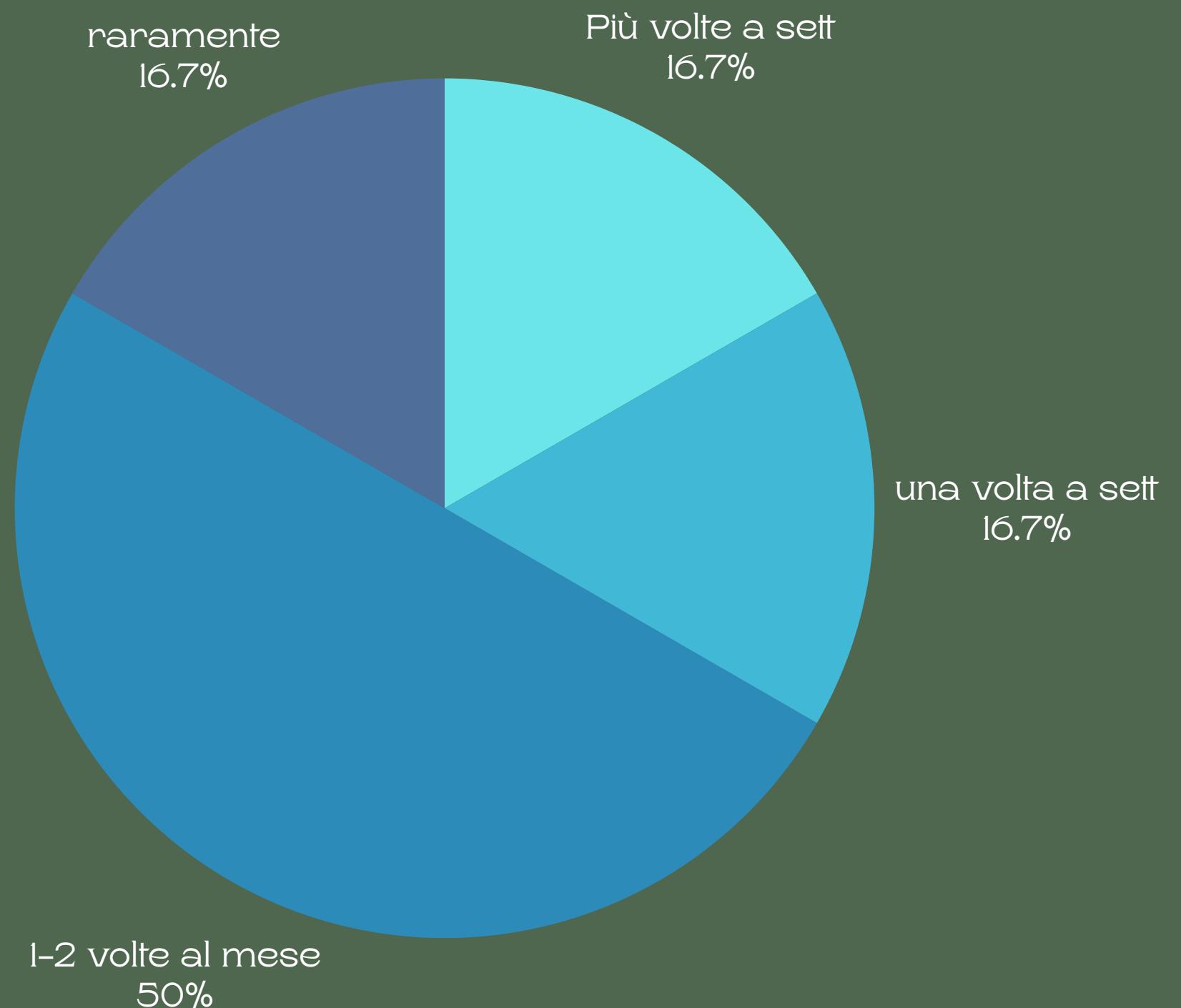
Età



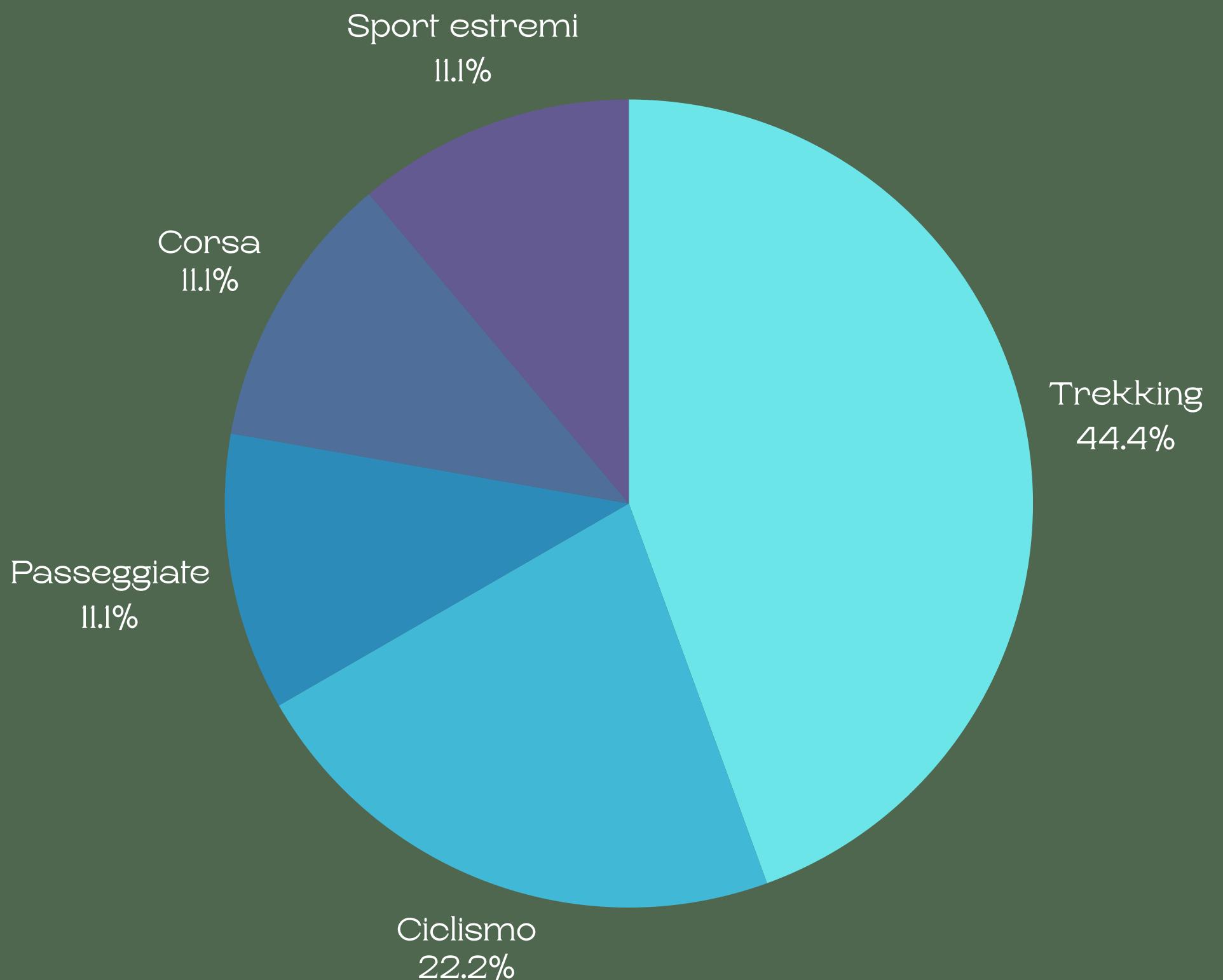
Occupazione



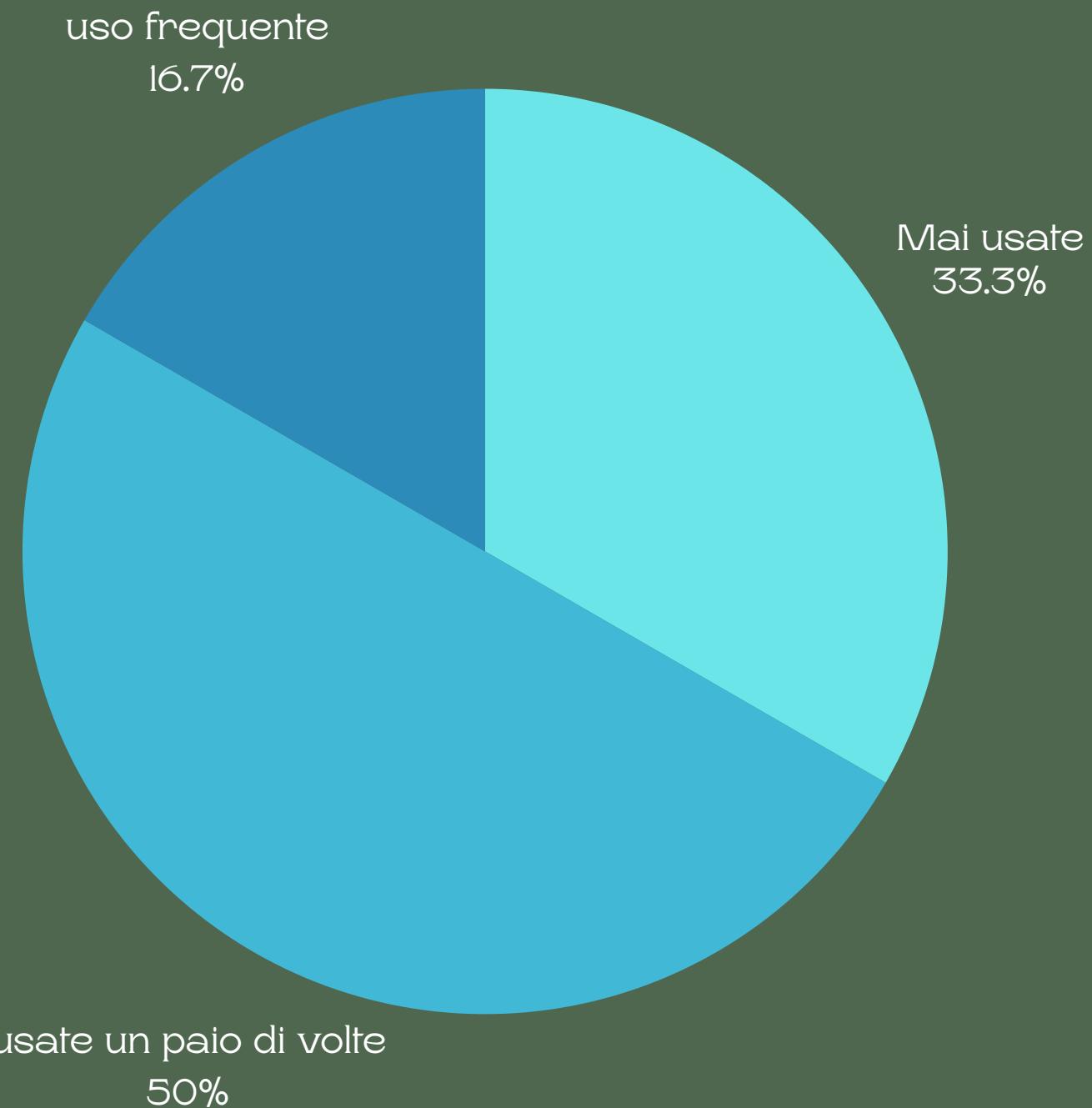
Quanto spesso vivi esperienze nella natura



Che tipo di esperienze prediliги



Quanta
familiarità hai
con app per
organizzare
eventi



Esecuzione dei task

Il test sarà condotto adottando la modalità THINK ALOUD, verrà quindi chiesto all'utente di esternare ad alta voce i propri pensieri e le proprie riflessioni durante lo svolgimento dei task.

TASK 1: *trovare un'attività*

Viene presentato all'utente il seguente scenario: è da tanto che vuoi fare un trekking in montagna, trova un'attività già avviata che faccia al caso tuo.

TASK 2: *creare una nuova attività*

Viene presentato all'utente il seguente scenario: non sei riuscito a trovare l'attività che avevi in mente, creane una da zero che soddisfi tutte le tue aspettative.

TASK 3: *modificare un'attività creata*

Viene presentato all'utente il seguente scenario: hai creato la tua attività, ma ti sei reso conto di aver sbagliato a compilare il campo mood, e invece che un'esperienza rilassante, te ne immaginavi una adrenalinica, fai questa modifica.

TASK 4: *disicriversi da un'attività*

Viene presentato all'utente il seguente scenario: ti sei ricordato che domani hai un appuntamento, controlla che non ti fossi iscritto ad un'attività per quel giorno, e in caso disiscriviti da essa.

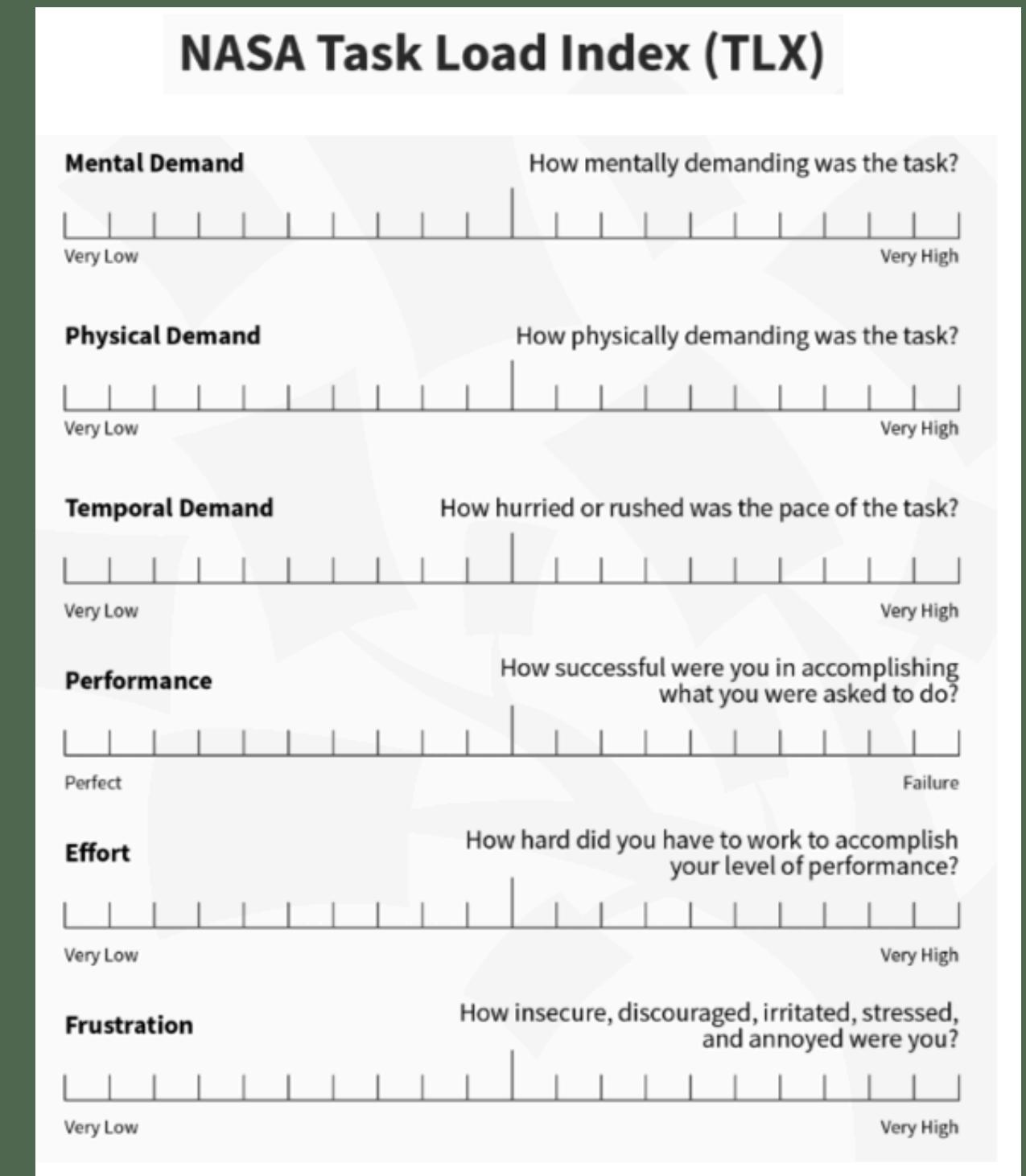
TASK 5: *valutare un'attività passata*

Viene presentato all'utente il seguente scenario: hai vissuto l'esperienza da te creata e ne sei pienamente soddisfatto, vai a recensirla.

Spiegazione delle metriche

Useremo le seguenti 3 metriche per valutare ogni task:

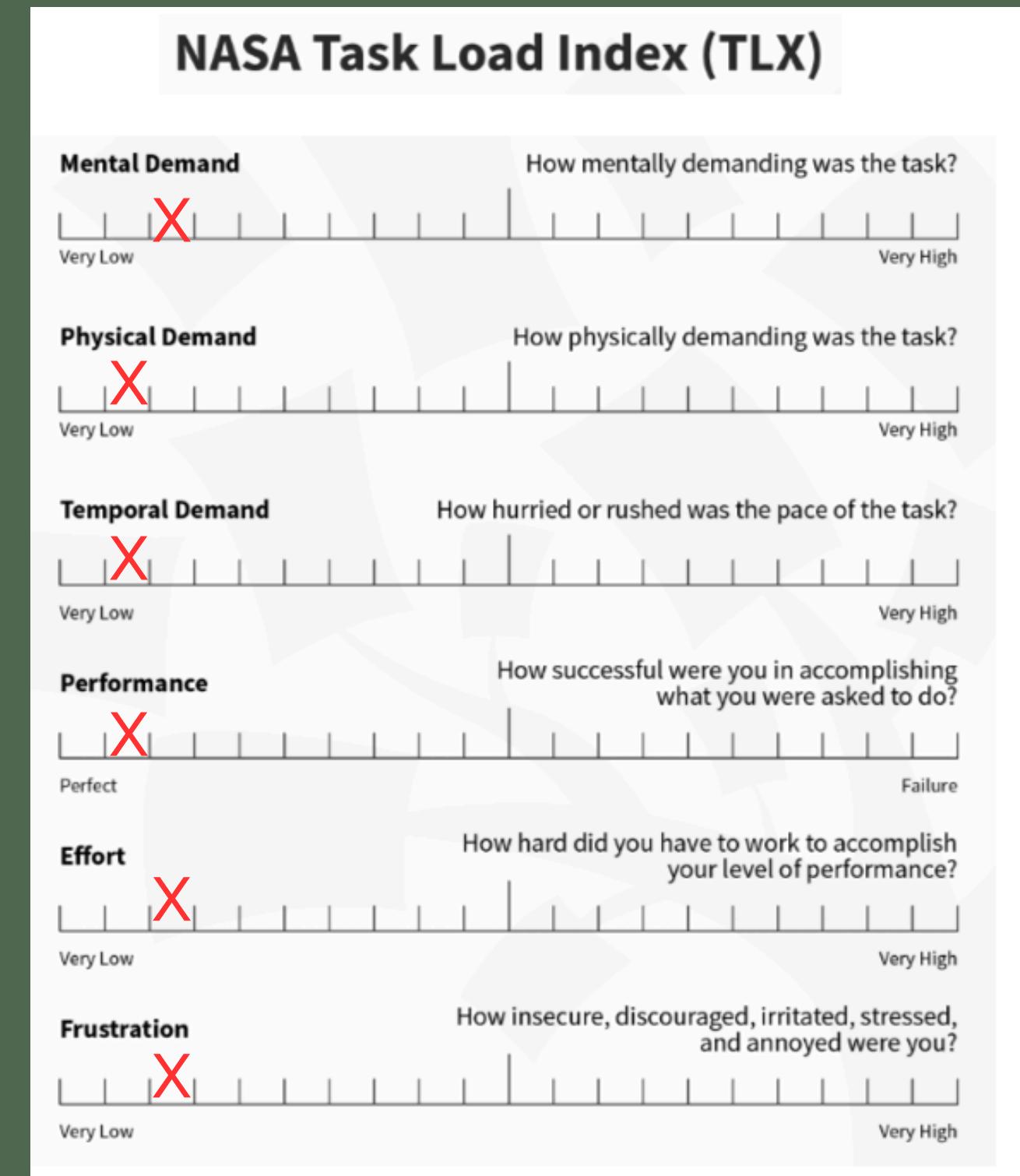
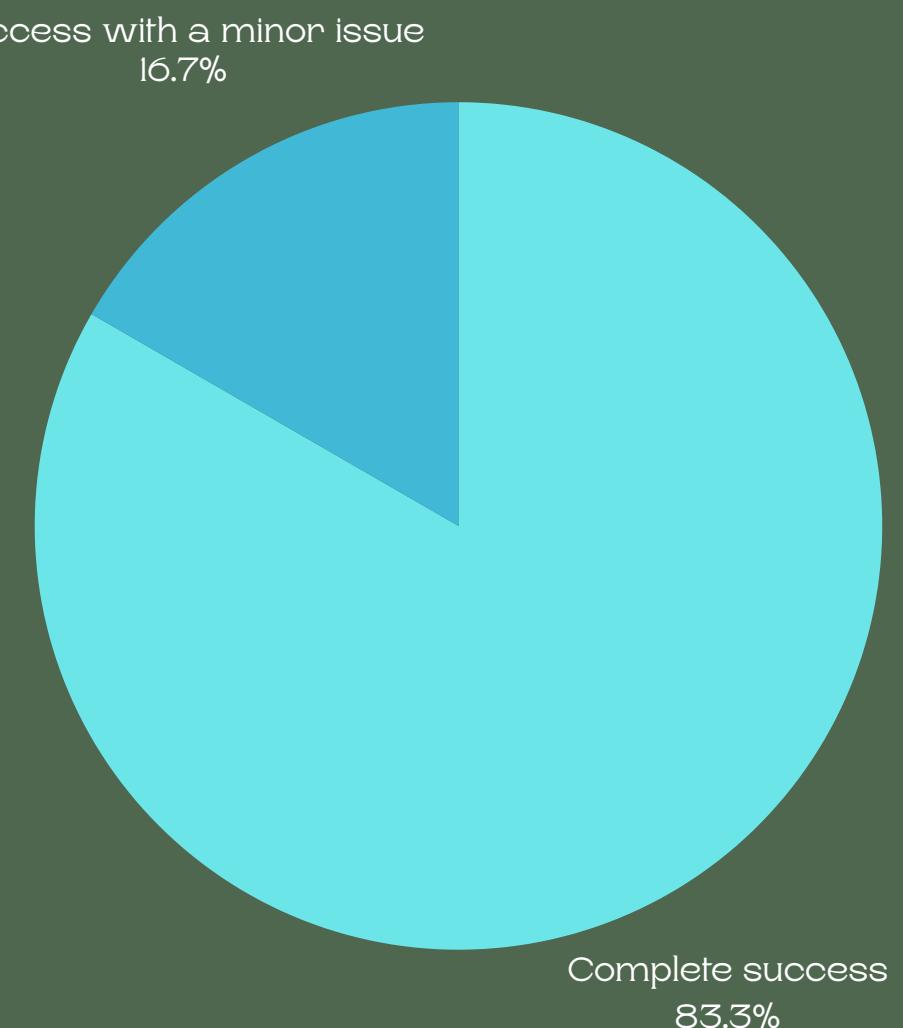
- media della distanza tra il tempo di esecuzione del partecipante e il tempo previsto (stimato da noi)
- tasso di successo del task valutato sulla seguente scala: complete success, success with a minor issue, success with a major issue, failure.
- media dei NASA TLX (foto a dx)



il campo “Physical Demand” è stato spiegato ai partecipanti col significato di numero di click

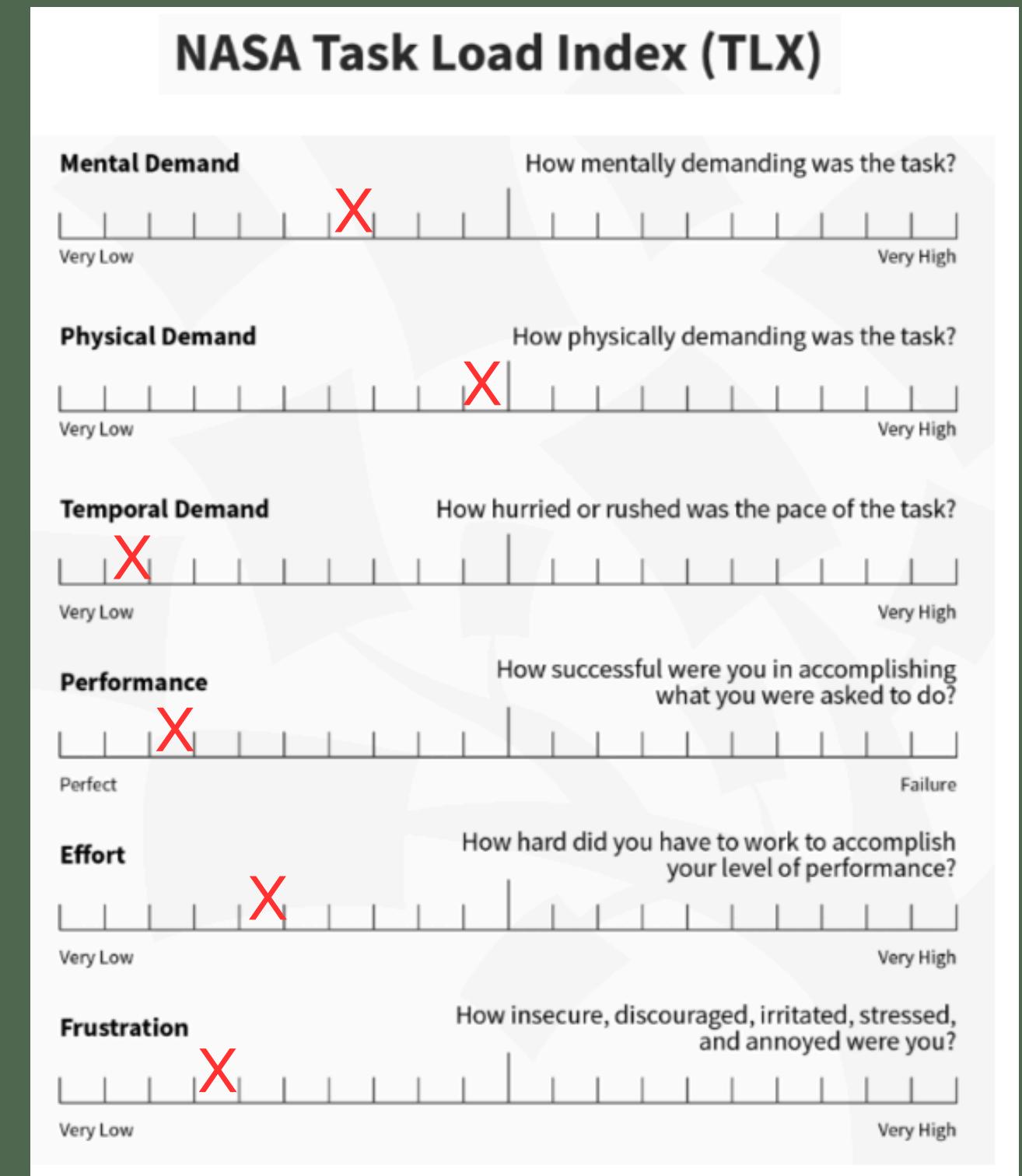
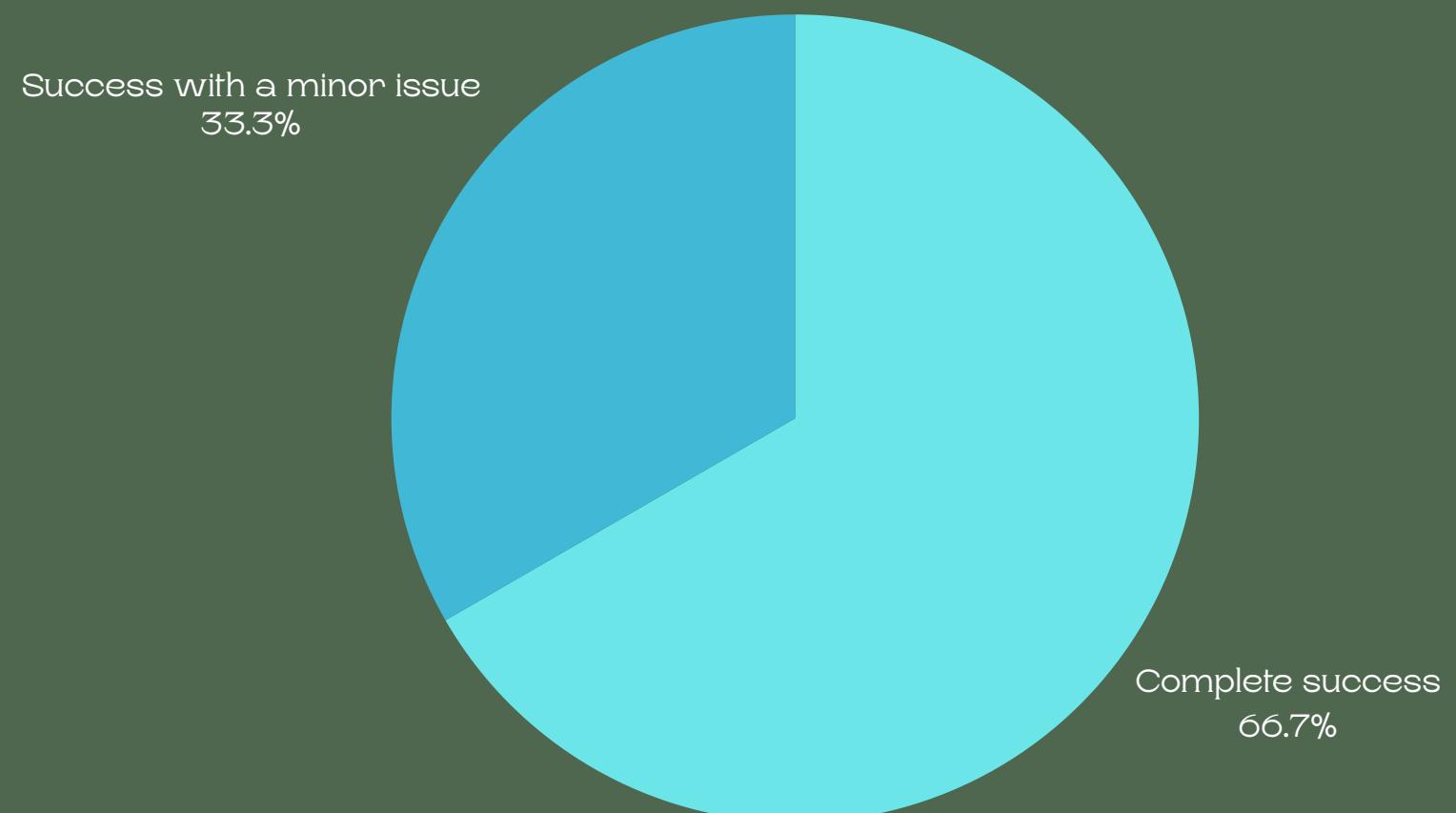
Task 1: trovare un'attività

tempo previsto: 5 sec
tempo medio di esecuzione: 8 sec



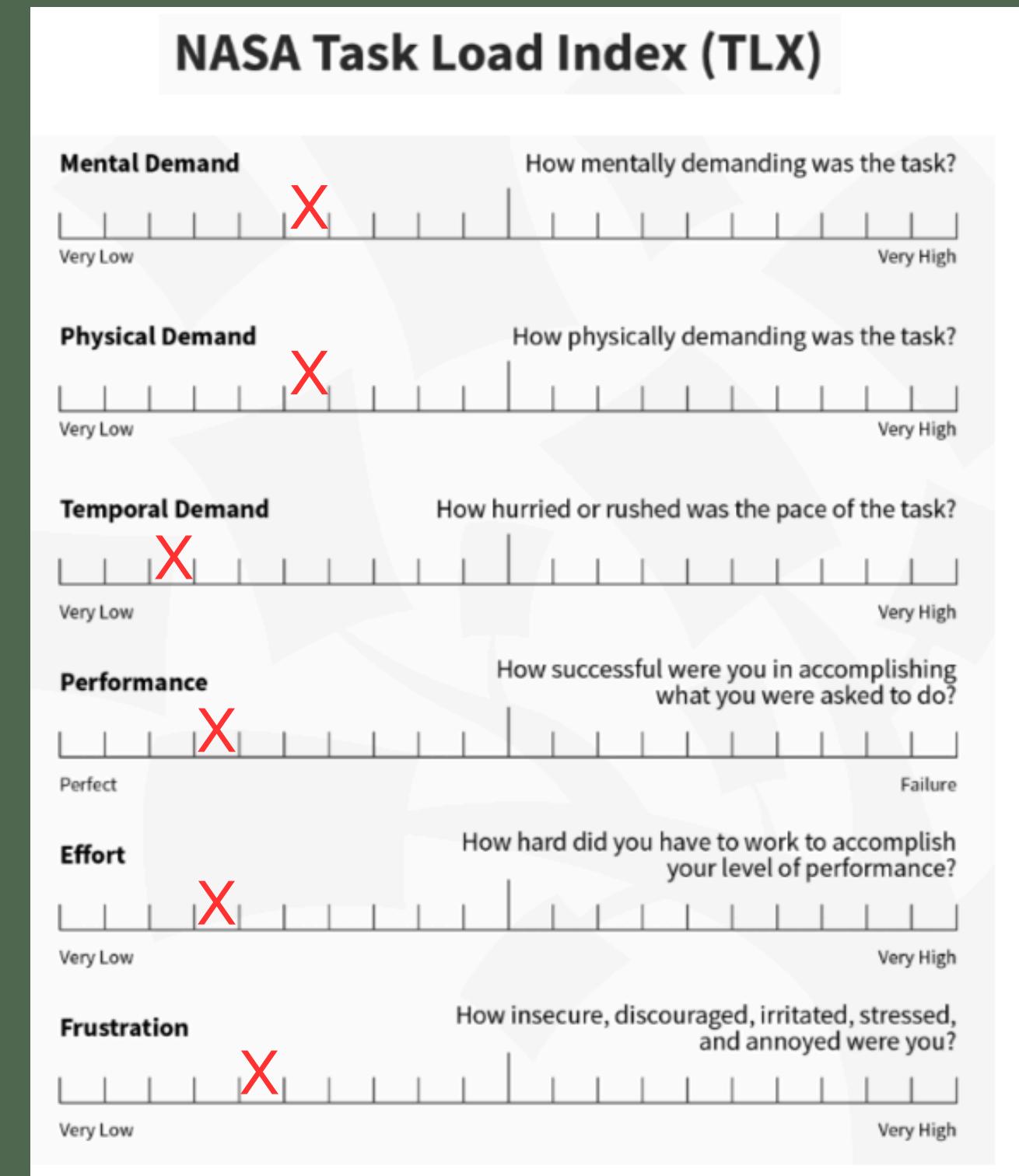
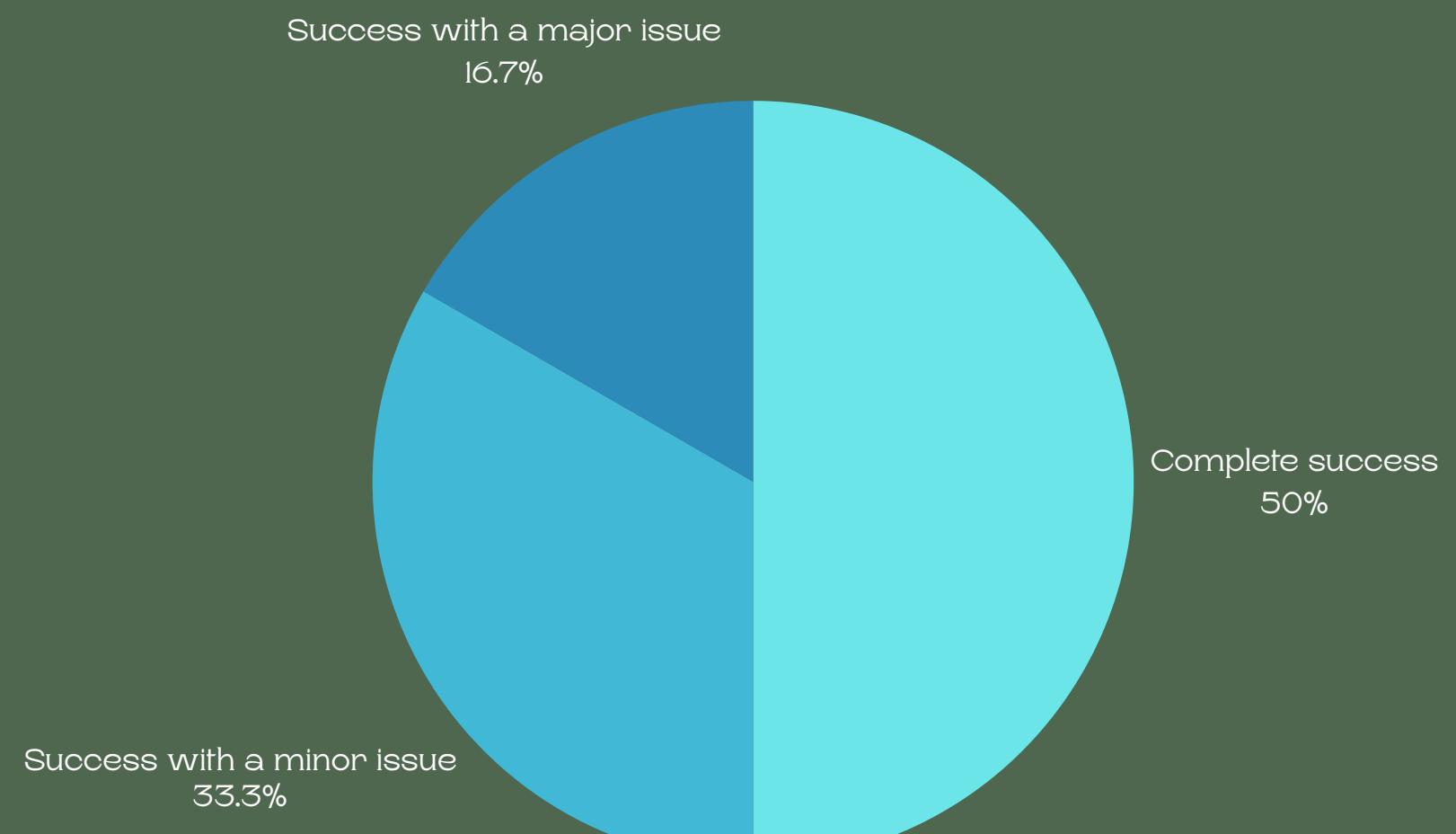
Task 2: creare una nuova attività

tempo previsto: 45 sec
tempo medio di esecuzione: 58 sec



Task 3: modificare un'attività creata

tempo previsto: 30 sec
tempo medio di esecuzione: 46 sec



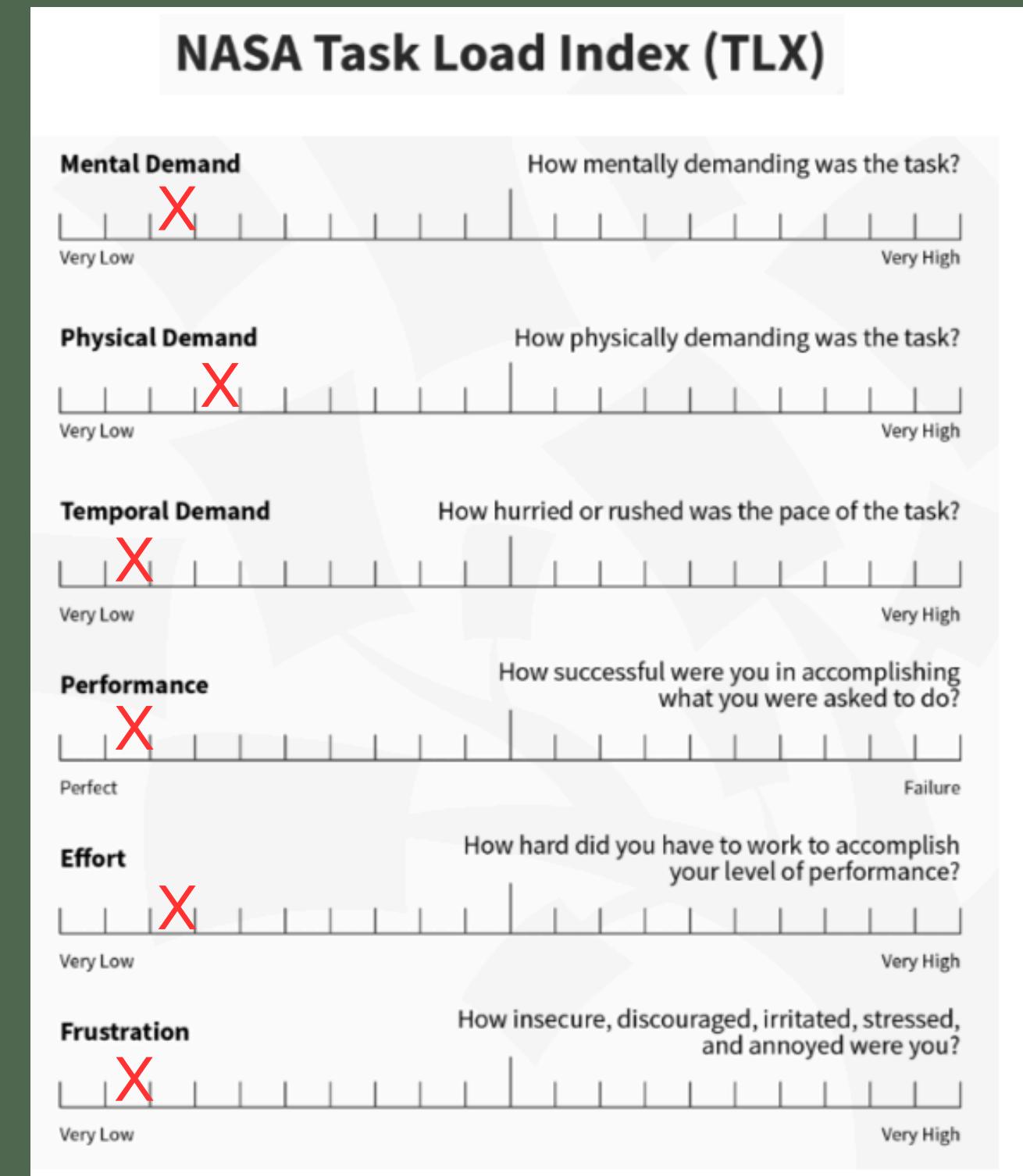
Task 4: disiscriversi da un'attività

tempo previsto: 10 sec

tempo medio di esecuzione: 8 sec



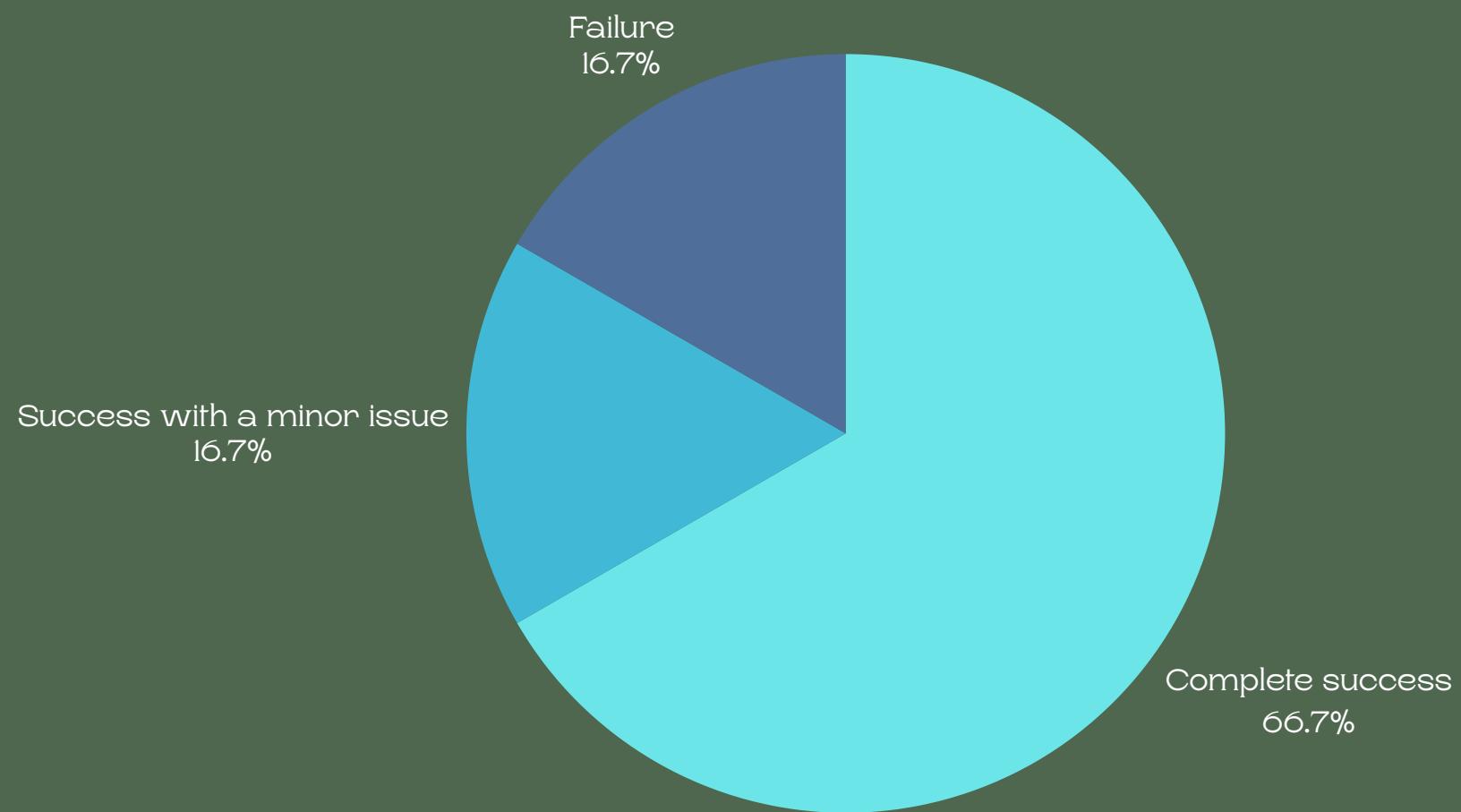
Complete success
100%



Task 5: valutare un'attività passata

tempo previsto: 30 sec

tempo medio di esecuzione: 54 sec



Il fallimento è stato dovuto al fatto che l'utente non ha trovato il pulsante valuta, che è stato successivamente riposizionato in modo che fosse maggiormente visibile.

NASA Task Load Index (TLX)

The NASA TLX is a questionnaire used to measure subjective mental workload. It consists of six subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Each subscale is rated on a scale from Very Low to Very High, with a midpoint labeled Average. A red 'X' marks the rating for each subscale.

Subscale	Question	Rating Scale
Mental Demand	How mentally demanding was the task?	Very Low — Average — Very High
Physical Demand	How physically demanding was the task?	Very Low — Average — Very High
Temporal Demand	How hurried or rushed was the pace of the task?	Very Low — Average — Very High
Performance	How successful were you in accomplishing what you were asked to do?	Perfect — Average — Failure
Effort	How hard did you have to work to accomplish your level of performance?	Very Low — Average — Very High
Frustration	How insecure, discouraged, irritated, stressed, and annoyed were you?	Very Low — Average — Very High

Questionario post-test

System Usability Scale (SUS)

Per ogni partecipante mostreremo il SUS e calcoleremo il corrispondente punteggio usando il rinomato algoritmo:

1. ogni risposta vale tra 1 e 5
2. per ogni domanda dispari, sottrai 1 al punteggio ($X-1$)
3. per ogni domanda pari, sottrai il punteggio da 5 ($5-X$)
4. somma i punteggi ottenuti
5. moltiplica il totale per 2,5

I think that I would like to use this product frequently.



I found the product unnecessarily complex.



I thought this product was easy to use.



I think that I would need the support of a technical person to be able to use this product.



I found the various functions in this product were well integrated.



I thought there was too much inconsistency in this product.



I would imagine that most people would learn to use this product very quickly.



I found this product very awkward to use.



I felt very confident using this product.



I needed to learn a lot of things before I could get going with this product.



Partecipante 1

Punteggi: 2 3 3 4 3 3 4 4 3 4

Somma: 33

Punteggio finale: 82,5

System Usability Scale (SUS)

Strongly Disagree

Strongly Agree

I think that I would like to use this product frequently.



I found the product unnecessarily complex.



I thought this product was easy to use.



I think that I would need the support of a technical person to be able to use this product.



I found the various functions in this product were well integrated.



I thought there was too much inconsistency in this product.



I would imagine that most people would learn to use this product very quickly.



I found this product very awkward to use.



I felt very confident using this product.



I needed to learn a lot of things before I could get going with this product.



Partecipante 2

Punteggi: 2 3 3 3 2 3 3 4 3 4

Somma: 30

Punteggio finale: 75

System Usability Scale (SUS)

Strongly Disagree

Strongly Agree

I think that I would like to use this product frequently.



I found the product unnecessarily complex.



I thought this product was easy to use.



I think that I would need the support of a technical person to be able to use this product.



I found the various functions in this product were well integrated.



I thought there was too much inconsistency in this product.



I would imagine that most people would learn to use this product very quickly.



I found this product very awkward to use.



I felt very confident using this product.



I needed to learn a lot of things before I could get going with this product.



Partecipante 3

Punteggi: 4 4 3 4 4 3 1 4 4 1

Somma: 32

Punteggio finale: 80

System Usability Scale (SUS)

Strongly Disagree

Strongly Agree

I think that I would like to use this product frequently.



I found the product unnecessarily complex.



I thought this product was easy to use.



I think that I would need the support of a technical person to be able to use this product.



I found the various functions in this product were well integrated.



I thought there was too much inconsistency in this product.



I would imagine that most people would learn to use this product very quickly.



I found this product very awkward to use.



I felt very confident using this product.



I needed to learn a lot of things before I could get going with this product.



Partecipante 4

Punteggi: 1 4 4 4 4 4 4 3 2 4

Somma: 34

Punteggio finale: 85

System Usability Scale (SUS)

Strongly Disagree

Strongly Agree

I think that I would like to use this product frequently.



I found the product unnecessarily complex.



I thought this product was easy to use.



I think that I would need the support of a technical person to be able to use this product.



I found the various functions in this product were well integrated.



I thought there was too much inconsistency in this product.



I would imagine that most people would learn to use this product very quickly.



I found this product very awkward to use.



I felt very confident using this product.



I needed to learn a lot of things before I could get going with this product.



Partecipante 5

Punteggi: 2 2 3 4 3 2 3 3 2 3

Somma: 27

Punteggio finale: 67,5

System Usability Scale (SUS)

Strongly Disagree

Strongly Agree

I think that I would like to use this product frequently.



I found the product unnecessarily complex.



I thought this product was easy to use.



I think that I would need the support of a technical person to be able to use this product.



I found the various functions in this product were well integrated.



I thought there was too much inconsistency in this product.



I would imagine that most people would learn to use this product very quickly.



I found this product very awkward to use.



I felt very confident using this product.



I needed to learn a lot of things before I could get going with this product.



Partecipante 6

Punteggi: 2 4 3 4 3 3 3 4 4 3 4

Somma: 34

Punteggio finale: 85

System Usability Scale (SUS)

Strongly Disagree

Strongly Agree

I think that I would like to use this product frequently.



I found the product unnecessarily complex.



I thought this product was easy to use.



I think that I would need the support of a technical person to be able to use this product.



I found the various functions in this product were well integrated.



I thought there was too much inconsistency in this product.



I would imagine that most people would learn to use this product very quickly.



I found this product very awkward to use.



I felt very confident using this product.



I needed to learn a lot of things before I could get going with this product.



Riepilogo

Abbiamo fatto la media dei punteggi finali ottenuti dai partecipanti, ed affidandoci alla tabella riportata sotto, è risultato che il nostro prototipo ha un Accettabile usabilità, con un punteggio SUS di 79

Punteggio SUS	Giudizio Estetico	Grado (Grade)	Usabilità
> 80.3	Eccellente (A+)	A	Alta Promozione
68 - 80.2	Buono (B)	B / C	Accettabile
68	Media Globale	C	Soglia minima
51 - 67	Sufficiente (D)	D	Marginale / Da migliorare
< 51	Scarso (F)	F	Inaccettabile / Critico

Considerazioni finali

Dall'esecuzione dei task è emerso che la maggior parte dei partecipanti si aspettava una maggiore implementazione delle funzionalità del sistema, quindi la nostra priorità futura sarebbe sicuramente quella di implementare funzionalità avanzate per rendere il sistema più completo.